International Journal of Engineering, Management and Humanities (IJEMH)
Volume 5, Issue 3, May.-June, 2024 pp: 112-118 ISSN: 2584-2145
www.ijemh.com

Mind Reading Model

Ekene F. Ozioko
Computer Science Department. Enugu State University of Science and Technology, Nigeria.
E-mail addresses: ekene.ozioko@esut.edu.ng (0000-0001-7092-6261),

Date of Submission: 15-05-2024 Date of Acceptance: 29-05-2024

ABSTRACT

People express their mental states all the time, even when interacting with machines. These mental states shape the decisions that we make, govern how we communicate with others, and affect our performance. The ability to attribute mental states to others from their behavior, and to use that knowledge to guide one's actions and predict those of others is known as the theory of mind or mind-reading. The principal contribution of this article is the real-time inference of a wide range of mental states from the head and facial displays in a video stream. In particular, the focus is on the premise of complex mental states: the effective and cognitive states of mind that are not part of the set of basic emotions. The automated mental state inference system is inspired by and draws on the fundamental role of mind-reading in communication and decisionmaking. The article describes the design, implementation, and validation of a computational model of mind-reading. The design is based on the results of several experiments that have been undertaken to analyze the facial signals and dynamics of complex mental states. The resulting model [1] is a multi-level probabilistic graphical model that represents the facial events classification in a raw video stream at different levels of spatial and temporal abstraction in accordance with the postulation from [2]. Dynamic Bayesian Networks model observable head and facial displays, and corresponding has hidden mental states over time from the work of [?], [3]. The automated mind-reading model is implemented by combining top-down predictions of mental state models with bottomup vision-based processing of the face. To support intelligent human-computer interaction, the system meets three important criteria. These are full automation so that no manual reprocessing or segmentation is required, real-time execution, and the categorization of mental states early enough after their onset to ensure that the resulting knowledge is current and valuable. The system is evaluated in terms of recognition accuracy, generalization, and real-time performance for six broad classes of complex mental states—agreeing, concentrating, disagreeing, interested, thinking, and unsure, on two different corpora. The system successfully classifies and generalizes to new examples of these classes with accuracy and speed that are comparable to that of human recognition. The research presented here significantly advances the developing ability of machines to infer cognitive-effective mental states in real-time from nonverbal and non-facial expressions of people. By developing a real-time system for the inference of a wide range of mental states beyond the basic emotions. The scope of humancomputer interaction scenarios in which this technology can be integrated has been widened to cover over 70% of everyday business applications including medicine, electronics, and elearning. The proposed models is an important step towards building socially and emotionally intelligent computers.

Keywords: Brain, Interface, Reading.

OVERVIEW

This research work draws inspiration from several disciplines of human mind interpretations. In the introductory section, the different theories on how humans perceive and interpret the mental and emotional states of others is presentated. This is followed by a review of the state of the art research done on how to enable computers to understand and mimic the human reseaning functionalities. The literature review started with the the basic human emotions, which have received most of the attention to date, and then investigated the mental states of human being. Section two is followed by the research methodology, discussion, conclusion/summary and lastly the direction of future research work.

I. INTRODUCTION

Mind-reading refers to the set of representational abilities that allow one to make inferences about others' mental states. In colloquial English, mind-reading is the act of "discerning, or appearing to discern, the thoughts of another person" or "guessing or knowing by intuition what somebody is thinking". Following the works of [4], [5], this dissertation uses mind-reading in a scientific sense to denote the set of abilities that allow a person to infer others' mental states from nonverbal cues and observed behaviour. From the point of view of an observer who mind-reads, the input is an array of observations, such as visual, auditory and even tactile stimuli, as well as context cues; the output is a set of mental states that are attributed to others. The types of mental states that people exhibit and attribute to each other include emotions, cognitive states, intentions, beliefs, desires and focus of attention. Mind-reading is often referred to in the developmental psychology literature as a specific faculty, separable from more general cognitive abilities such as general intelligence and executive function. Interest in the functions and mechanisms of this ability has become a central and compelling question for cognitive scientists in recent years. Since Premack and Woodruff and Dennett first stimulated the interest of cognitive scientists in mind-reading, numerous tasks, methods and theories have accumulated in the literature on this topic. Developmental and experimental studies, as in Goldman and Sirpada, investigate theoretical models of how people mind-read. Other studies examine the neural basis of mind-reading using brain-imaging technologies like functional Magnetic Resonance Imaging. Examples include the studies by [6]–[8]. The findings from the two classes of studies contribute to our understanding of how the cognitive skills that enable high-level social cognition are organized in the human brain, and the role they play in everyday functioning. These findings also form the basis for the computational model of mind-reading in Chapter 2.

1.2.1 THE FUNCTIONS OF MIND-READING

While subtle and somewhat elusive, mind-reading is fundamental to the social functions we take for granted. It is an important component of a broader set of abilities referred to as social intelligence. Through mind-reading we are able to make sense of other people's behaviour and predict their future actions. It also allows us to communicate effectively with other people. In addition, mind-reading has been described as a cognitive component of empathy. A good empathizer can immediately sense when an emotional change has occurred in someone, what the causes of this change might be, and what might make this person feel better. Mind-reading is also a powerful tool in persuasion and negotiation: by realizing that people's thoughts and beliefs are shaped by the information to which they are exposed, it is possible to persuade them to change what they know or how they think. Mind-reading is also a key component of other processes such as perception, learning, attention, memory and decision-making. In their studies, LeDoux, Damasio and Adolphs uncover the parts of the brain that are responsible for higher order processing of emotion. These studies and others, like that by Purves, have shown that these brain areas are interconnected to other brain structures that are involved in the selection and initiation of future behaviour. These findings emphasize the interplay of emotion and cognition, and have led to a new understanding of the human brain, in which it is no longer considered as a purely cognitive information processing system; instead it is seen as a system in which affective and cognitive functions are inextricably integrated with one another. The implications for user-modelling in humancomputer interaction (HCI) are clear: an accurate model of the user would have to incorporate the affective as well as the cognitive processes that drive the user's reasoning and actions.

1.2.2 MIND-READING MECHANISMS

Mind-reading involves two components that originate in different parts of the brain and develop at distinctive ages. These components may be impaired selectively across different populations of people. The first component encompasses the social-perceptual component of mindreading, which involves detecting or decoding others' mental states based on immediately available, observable information. According to [9], [10], one could attribute the mental state confused to a person given their facial expressions and/or tone of voice. As its name implies, this component involves perceptual, or bottom-up processing of facial or other stimuli. It also involves cognitive abilities, or top-down processing of abstract models that depict how people's behaviour generally map to corresponding mental states. The second component is the social-cognitive component of mind-reading. This involves reasoning about mental states with the goal of explaining or predicting a person's actions. Examples include distinguishing jokes from lies, or predicting peoples' behaviour on the basis of false beliefs. False belief tasks test a person's understanding that other people's thoughts can be different from one another and from reality, and are the prototypical measure of the social-cognitive aspect of mind-reading. It is important to note that both the social-perceptual and the social-cognitive components of mind-reading are inherently uncertain—we are never 100% sure of a person's mental state. A person's mental state (John is thinking), and its content (what John is thinking about) are not directly available to an observer; instead they are inferred from observable behaviour and contextual information with varying degrees of certainty. Moreover, people often have expressions that reflect emotions or mental states that are different

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 5, Issue 3, May.-June, 2024 pp: 112-118 ISSN: 2584-2145
www.ijemh.com

than their true feelings or thoughts. The discrepancy between expressed and true feelings, such as in lying and deception, can sometimes be identified from fleeting, subtle micro-expressions. The problem of identifying deception from facial expressions is beyond the scope of this dissertation.

1.3 READING THE MIND IN THE FACE

Facial expressions are an important channel of nonverbal communication according to [?]. They communicate a wide range of mental states, such as those in Figure 1.1. Besides conveying emotions, facial expressions act as social signals that enhance conversations and regulate turn-taking. A face is comprised of permanent facial features that we perceive as components of the face such as the mouth, eyes and eyebrows, and transient features such as wrinkles and furrows. Facial muscles drive the motion and appearance of permanent facial features and produce transient wrinkles and furrows that we perceive as facial expressions from [?]. Head orientation, head gestures and eye gaze have also been acknowledged as significant cues in social-perceptual understanding. For example, Haidt et al. show that gaze aversion, a controlled smile and a head turn are signals of embarrassment. Langton et al. emphasize the role of head orientation and eye gaze as an indicator of the focus of attention. In 1971, Ekman and Friesan demonstrated the universal recognition of six emotions from the face in a number of cultures. The six emotions— happiness, sadness, anger, fear, surprise and disgust—became known as the basic emotions. From [?], the facial expressions associated with these basic emotions have almost dominated the study of facial expressions for the past forty years. These six emotions are viewed as dedicated neural circuits that facilitate adaptive responses to the opportunities and threats faced by a creature. For example, the feeling of fear leads to flight, while that of anger leads to fight. In addition to their universality, these emotions are also recognized by very young normally developing children.

Fig. 1: Peak frames for each of the six basic emotions, from L – R: ANGER, DISGUST, FEAR, JOY, SORROW, SURPRISE.

Our everyday social experiences, however, involve much more than just these six emotions, and the ability to recognize them needs to be studied. Rozin and Cohen describe a study in which college students were instructed to observe the facial expressions of other students in a university environment and to report the emotion being expressed. The most common facial expressions reported were those of confusion, concentration and worry. Despite their prevalence in everyday interactions, these facial expressions have not been investigated because they do not correspond to generally recognized emotions, leading the authors of the study to call for more studies that explore the facial expressions of mental states that are not typically thought of as emotions. Simon BaronCohen and his group at the Autism Research Centre at the University of Cambridge, have undertaken a series of studies to investigate the facial expressions of mental states other than the basic emotions. The principal objective of these studies is to investigate the differences in emotion processing between a general population of people and those with ASD. Because these differences were not apparent on basic emotion recognition tasks, [11] yet were clearly demonstrated in natural interaction contexts, more challenging tasks were needed. Baron-Cohen and Cross show that normally developing fouryear-old children can recognize when someone else is thinking from the direction of that person’s gaze. That is, when a person’s eyes are directed away from the viewer, to the left or right upper quadrant, and when there is no apparent object to which their gaze is directed, we recognize them as thinking about something. In Baron-Cohen et al., the cross-cultural recognition of paintings and drawings of the face was shown among normal adults and children for mental states such as scheme, revenge, guilt, recognize, threaten, regret and distrust. In two other studies, Baron-Cohen et al. show that a range of mental states, cognitive ones included, can be inferred from the eyes and the face. Figure 1.2 shows several examples of the face stimuli of complex mental states used in Baron-Cohen et al. The findings of these studies show that many mental states are like virtual print-outs of internal experience, simply waiting to be read by an observer (with a concept of mind).

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 5, Issue 3, May.-June, 2024 pp: 112-118 ISSN: 2584-2145
www.ijemh.com

Fig. 2: Four examples of the complex mental states face stimuli used in Baron-Cohen et al. [BWJ97]: (from left to right) GUILT vs. Arrogant; (b) THOUGHTFUL vs. Arrogant; (c) FLIRTING vs. Happy; (d) ARROGANT vs. Guilt. The correct responses are shown as uppercase letters.

II. COMPUTATIONAL MODEL OF MIND-READING

A person's mental state is not directly available to an observer; instead it is inferred from nonverbal cues such as facial expressions. Presented here is a novel approach to mental state representation based on the theory of mindreading. This approach combines vision-based perceptual processing with top-down reasoning to map low-level observable behaviour into high-level mental states.

2.1 REPRESENTATION OF MENTAL STATES

As shown in Fig. 1, we use Dynamic Bayesian Networks (DBNs) [12] to model the unfolding of mental states over time P (X[t]), where X is a vector of events corresponding to different mental states. A DBN is a graph that represents the causal probability and conditional independence relations among events that evolve over time. The hidden state of each DBN represents an event with two possible outcomes: true whenever the user is experiencing a specific mental state, and false otherwise. The observations or evidence nodes represent the recognized head and facial displays Y.

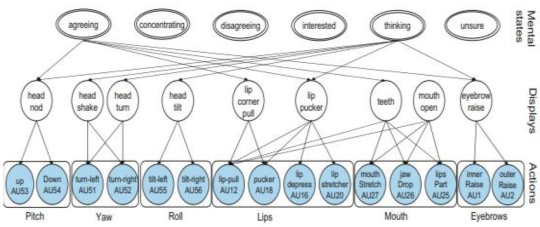

Fig. 3: Multi-level computational model of mind-reading. For clarity, the displays for only two mental states are shown.

The double circle around a mental state node encodes the temporal dependency between that node in consecutive slices of the network, Xi[t - 1] and Xi[t]. Having a model for each class means that the hidden state of more than one DBN can be true, so that co-occurring mental states can be represented by the system. The DBN parameters and structure are learnt from exemplar videos using maximum likelihood estimation and feature selection.

2.2 Observational Evidence: Head and Facial Displays

The observational evidence consists of the head and facial displays that

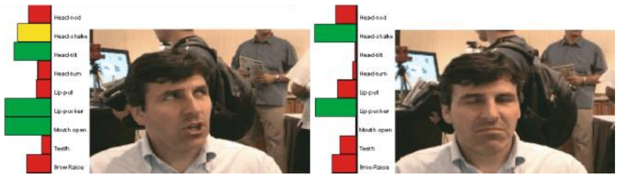

Fig. 4: Real time display recognition (frames sampled every 0.7s).

The bars represent the output probabilities of the HMM classifiers (top to bottom): head nod, shake, tilt, turn, lip corner pull, lip pucker, mouth open, teeth and eye-brow raise.

2.3 INFERENCE FRAMEWORK

Inference involves recursively updating the belief state of hidden states based upon the knowledge captured in the DBNs and available evidence—the head and facial displays that are recognized throughout a video, their dynamics (duration, relationship to each other, and when in the video they occur) and previous mental state inferences. We implement the inference framework as a sliding window of evidence (Algorithm 1).

At any instant t, the observation vector that is input to the inference engine is a vector of the w most-recent displays Y[t w : t], and the corresponding most-recent mental state inferences P*X*[*tw* : *t*1]. The output is a probability that the observation vector was generated by each of the DBNs. The inference engine uses the unrolled-junction-tree algorithm [14].

Algorithm 1. Mental state inference

Objective: $P(X_i[t])$, the belief state of $1 \le i \le x$ mental states over time $1 \le t \le T$
Given: x DBNs with y observations nodes; evidence length w and sliding factor dw
Instantiate inference engine
for all t in w time slices **do**
Get current observations $\mathbf{Y}[t]$
for all t in T time slices **do**
Enter evidence so far: $\mathbf{Y}[t-w:t]$ and $P(\mathbf{X}[t-w:t-1])$
Calculate marginal probabilities $P(\mathbf{X}[t])$
Advance window $t = t + dw$
Get current observations $\mathbf{Y}[t]$

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 5, Issue 3, May.-June, 2024 pp: 112-118 ISSN: 2584-2145
www.ijemh.com

III. EXPERIMENTAL EVALUATION

[13], [14], trained and tested their system on videos from the Mind-Reading DVD (MR) [2], a guide to emotions developed for Autism Spectrum Disorders. An upper bound of 88.9% and an average accuracy of 77.4% was achieved for agreeing, concentrating, disagreeing, interested, thinking and unsure. To test if the system generalized beyond the controlled videos in MR, they collected videos at the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2004). Fig. 5 (left) shows frames of both corpora.

3.1 The CVPR 2004 Corpus

They asked 16 conference attendees to act six mental states: agreeing, concentrating, disagreeing, interested, thinking and unsure. The volunteers were not given any instructions on how to act the mental states, which resulted in considerable within-class variation between the 16 videos of each emotion. They were asked to name the mental state they would act immediately before they started; this was later used to label the videos. Unlike prevalent facial expression databases [10], they placed no restrictions on the head or body movements of volunteers. All 16 volunteers were aged between 16 and 60 and worked in computer-science or engineering; most were males of a white ethnic origin. Their characteristics are summarized in Table 1. The videos were captured at 30 fps at a resolution of 320x240 and were labelled using the audio accompanying the footage. The background of the videos was dynamic: people were moving in and out of the neighbouring demonstration booth. They just relied on the lighting in the conference room at the time. The face-size varies within and between videos as the volunteers moved toward/away from the camera. By contrast, the actors in MR had a frontal pose at a constant distance from the camera, none wore glasses or had facial hair and the videos all had a uniform white background and the lighting was professionally set up. Eight videos (15s) were discarded: three lasted less than two seconds which is when the first DBN invocation occurs, and the system failed to locate the face in five videos. They used the remaining 88 videos (313s).

Table 1. Characteristics of CVPR corpus "actors". Gender: Male • Female ○; Ethnicity: White • Asian ○; Glasses • Facial hair ○; Looking down • Talking ○.

Subject ID	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Gender	•	•	○	•	•	•	•	•	○	•	○	•	•	•	•	•
Ethnicity	•	•	•	•	•	○	•	•	•	•	•	•	•	•	•	○
Glasses/Facial hair	○				•	•					• ○					
Frontal	•	•	•	•	•	•	•		•	•	•	•	•	•	•	•
Looking down/Talking		○	○	•				○	•	○						

3.2 Human Baseline

Having been posed by people who were not professional actors, the CVPR videos were likely to include incorrect or bad examples of a mental state, and were weakly labelled. To establish a baseline with which to compare the results of the system, they tested how a panel of people would classify the videos. A forced-choice procedure was adopted, with six choices on each question: agreeing, concentrating, disagreeing, interested, thinking, unsure. Chance responding was 16.7%. Participants were shown a video on a projection screen, and then asked to circle only one mental state word that best matched what the person in the video was feeling. The panel consisted of 18 participants (50.0% male, 50.0% female), mostly software developers between the ages of 19 and 28. The test generated 88 trials per participant for a total of 1584 responses. The distribution of results is shown in Fig. 5 (right). The percentage of correct answers ranged from 31.8% to 63.6% (mean=53.03%, SD=0.068). The agreement-score of a video—the percentage of panel participants who assigned the same label to a video—varied between 0- 100%. Only 11% of the videos achieved an agreement-score of 85% or more on the truth label of the video; these were deemed as good examples of mental states. The confusion matrix of responses is shown in Fig. 5 (left). The classification rate is highest for disagreeing (77.5%) and lowest for thinking (40.1%). For a false positive rate of 9.4%, the recognition accuracy of the panel was 54.5%.

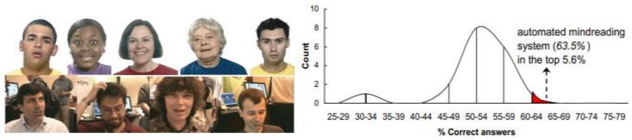

Fig. 5: (top-left) Mind Reading DVD; (bottom-left) CVPR corpus; (right) Distribution of human responses. The accuracy of the system is also shown.

3.3 Results of Computational Model of Mind-Reading

[13]–[15] trained the system on MR videos and tested it on the 88 videos of the CVPR corpus. A classification is correct if the mental state scoring the minimum error (i.e. largest area under the curve) matches the ground-truth label of the video. Fig. 4 shows an example of a 4.3-second long video labelled as thinking (77.8% agreement-score). A (false) head shake, a head tilt, a head turn and a lip-pull were recognized. Since thinking spans the largest area and this matches the groundtruth label of the video, this is a correct classification. The results are summarized in Fig. 5 (right). The classification rate is highest for

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 5, Issue 3, May.-June, 2024 pp: 112-118 ISSN: 2584-2145
www.ijemh.com

disagreeing (85.7%) and lowest for thinking (26.7%)—all higher than chance responding (16.7%). For a mean false positive rate of 7.3%, the overall accuracy of the system is 63.5%. Compared with the results of humans classifying the exact set of videos, the automated mind-reading system scores among the top 5.6% of humans, and 10.2% better than the mean accuracy reported in the sample of 18 people. The result is superimposed on the distribution of human responses shown in Fig. 3 (right). The principal reason why both human recognition (54.5%) and the system's accuracy (63.5%) is generally low is the untrained acting and weak labelling of the CVPR corpus videos. In addition, the recording conditions of the CVPR corpus were much less controlled than that of MR, resulting in challenges in processing these videos automatically (e.g., speech and changes in lighting conditions).

According to [?], the system's recognition accuracy increases to 80% for the 11% of videos with agreement-score of

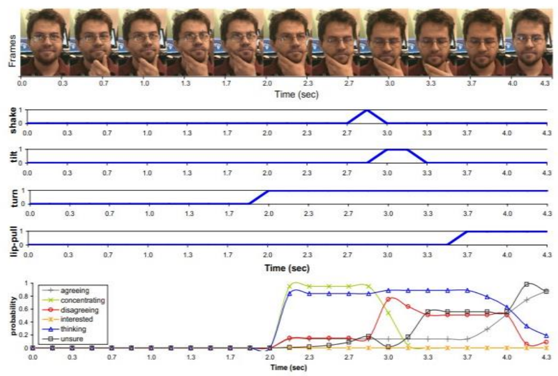

Fig. 4. Mental state inference: (top) frames from a video labelled as *thinking* (CVPR Corpus); (middle) head and facial displays; (bottom) mental state inferences

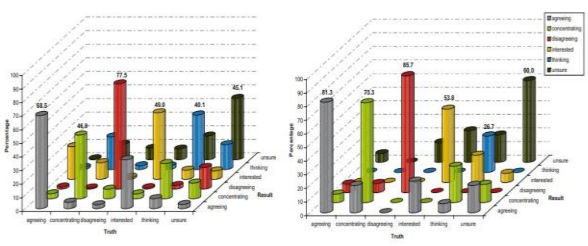

Fig. 5. Confusion matrix of results shown as a 3D bar chart: (left) human recognition results of the CVPR corpus; (right) the system's recognition the CVPR corpus

85% or more, a result similar to that obtained from evaluating the system on MR.

IV. CONCLUSION

Existing human-computer interfaces are mindblind—oblivious to the user's mental states and intentions. These user interfaces have zero persuasive power, cannot initiate interactions with the user, and are mostly limited to a command and control interaction paradigm. Even if they do take the initiative, like the now retired Microsoft Clip, they are often misguided and irrelevant, and end up frustrating the user. With the increasing complexity of HCI and the ubiquity of mobile and wearable devices, a new interaction paradigm is needed in which systems autonomously gather information about the user's mental state, intentions and surrounding context to adaptively respond to that. In this dissertation I have described the design and implementation of a real time system for the inference of complex mental states from head and facial signals in a video stream. The computational model of mind-reading presents a coherent framework for incorporating mind-reading functions in user interfaces. The implementation of the system has shown that it is possible to infer a wide range of complex mental states from the head and facial displays of people, and that it is possible to do so in real time and with minimal lag. Moving forward, there are numerous research opportunities that warrant further research. The computational model of mind-reading can be extended to more modalities and context cues in order to recognize a wider range of mental states. A more rigourous learning mechanism needs to be implemented that fuses these different sensors in an efficient way. The model needs to generalize well to naturally evoked mental states, and applications of automated mind-reading in HCI need to be conceptualized, implemented and validated. As the challenges presented in this dissertation are addressed over time, information about a user's mental state will become as readily available to computer applications as are keyboard, mouse, speech and video input today. Interaction designers will have at their disposal a powerful new tool that will open up intriguing possibilities not only in verticals such as assistive technologies and learning tools, but also in applications we use in our day-to-day lives to browse the web, read emails or write documents. The result will be next-generation applications that employ the user's emotional state to enrich and enhance the quality of interaction, a development that will undoubtedly raise the complexity of human-computer interactions to include concepts such as exaggeration, disguise and deception that were previously limited to human-to-human interaction. The research presented here serves as an important step towards achieving this vision. By developing a computational model of mind-reading that infers complex mental states in real time, the scope of human-computer interaction scenarios in which automated facial analysis systems can be integrated has been widened. It also motivates future research that takes full advantage of the rich modality of the human face and of nonverbal cues in general, to further the development of socially and emotionally intelligent interfaces.

a) Funding Statement: This work was supported by the Nigerian Tertiary Education trust Fund in collaboration with Enugu state University of Science and Technology [TEDFund scholarship grant 2017]

REFERENCES

[1]. A. Kapoor, R. W. Picard, and Y. Ivanov, "Probabilistic combination of multiple modalities to detect interest," in Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., vol. 3. IEEE, 2004, pp. 969–972.
[2]. H. Gu and Q. Ji, "Facial event classification with task oriented dynamic bayesian network," in Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004., vol. 2. IEEE, 2004, pp. II–II.
[3]. G. Littlewort, M. S. Bartlett, I. Fasel, J. Susskind, and J. Movellan, "Dynamics of facial expression extracted automatically from video," in 2004 Conference on Computer Vision and Pattern Recognition Workshop. IEEE, 2004, pp. 80–80.
[4]. S. Baron-Cohen, "Mindblindness: An essay on autism and theory of mind. cambridge, massachussets: Bradford book," 1995.
[5]. S. Baron-Cohen, O. Golan, S. Wheelwright, and J. Hill, "Mind reading: The interactive guide to emotions. london: Jessica kingsley," 2004.
[6]. S. Baron-Cohen, S. Wheelwright, J. Hill, Y. Raste, and I. Plumb, "The "reading the mind in the eyes" test revised version: a study with normal adults, and adults with asperger syndrome or high-functioning autism," The Journal of Child Psychology and Psychiatry and Allied Disciplines, vol. 42, no. 2, pp. 241–251, 2001.
[7]. B.-C. Simon, W. Sally, H. Jacqueline, R. Yogini, and P. Ian, "The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism," Journal of child psychology and psychiatry, vol. 42, no. 2, pp. 241–251, 2001.
[8]. P. Ekman and W. V. Friesen, "Facial action coding system," Environmental Psychology & Nonverbal Behavior, 1978.
[9]. B. Fasel and J. Luettin, "Automatic facial expression analysis: a survey," Pattern recognition, vol. 36, no. 1, pp. 259–275, 2003.
[10]. M. Pantic and L. J. M. Rothkrantz, "Automatic analysis of facial expressions: The state of the art," IEEE Transactions on pattern analysis and machine intelligence, vol. 22, no. 12, pp. 1424–1445, 2000.
[11]. M. Pardas, A. Bonafonte, and J. L. Landabaso, "Emotion recognition based on mpeg-4 facial animation parameters," in 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 4. IEEE, 2002, pp. IV–3624.
[12]. K. P. Murphy, Dynamic bayesian networks: representation, inference and learning. University of California, Berkeley, 2002.
[13]. R. A. El Kaliouby, "Mind-reading machines: automated inference of complex mental states," University of Cambridge, Computer Laboratory, Tech. Rep., 2005.
[14]. R. e. Kaliouby and P. Robinson, "Real-time inference of complex mental states from facial expressions and head gestures," in Real-time vision for human-computer interaction. Springer, 2005, pp. 181–200.
[15]. M. H. Mahoor, S. Cadavid, D. S. Messinger, and J. F. Cohn, "A framework for automated measurement of the intensity of non-posed facial action units," in 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. IEEE, 2009, pp. 74–80.